

# A STATISTICAL MODEL OF SEA CLUTTER IN PANCHROMATIC HIGH RESOLUTION IMAGES

*G. Jubelin, A. Khenchaf*

LABSTIC UMR CNRS 6285

Ecole Nationale Supérieure de Techniques Avancées Bretagne

2 rue François Verny, 29806 Cedex 9 Brest, France

## ABSTRACT

From the perspective of developing a ship detection algorithm on optical imagery, a statistical model is developed to approximate histograms from high-resolution images of the sea surface. This model is developed using an empirical approach based on analysis of hundreds of images acquired on all the oceans of the planet. Several statistical distributions are selected in agreement with the state of the art in remote sensing of sea surface and ship detection. Thumbnails of different sizes are extracted from satellite images, their histograms are then calculated. The generated histograms are approximated by the probability density functions of the different statistical distributions selected. The least-squares method is used. Reliability of the models is tested by applying the Kolmogorov-Smirnov test and analyzing the sum of squared residuals in least-squares sense. Alpha-stable distribution is retained as the best among tested models. Texture and frequency descriptors are calculated and compared with Alpha-stable parameters to assess relations binders. Reliability of the models according to the sensors, the sea state is discussed.

**Index Terms**— Sea clutter, optic, high resolution imagery, Alpha-stable distribution

## 1. INTRODUCTION

Statistical modeling of sea clutter in the optical images is very little studied. Reflectance of the sea surface as a function of environmental and viewing conditions is well known especially in the field of water color for estimation of the chlorophyll-a concentration for example. Statistical modeling of sea clutter is against highly developed for ship detection in synthetic aperture radar (SAR) images. This modeling has fueled a lot of literature that will be the basis of this study. Few studies on optical ship detection are based on Bayesian methods and so exploiting the sea clutter modeling. Among them Proia et al [1] used a Gaussian function to model both

the signature of the ship than the sea. There is a lack here that we will try to fill in this study.

A large amount of data is available for this study, 2790 mainly from optical images of satellite sensors SPOT4, SPOT5, KOMPSAT2, FORMOSAT2 provided by ASTRUM GEO-Information Services. They cover all the oceans of the planet and show a great variability of the sea state. We have defined a model as generic as possible to best approximate histograms whatever the sea state, the atmospheric conditions, and the sensor used. We work without a priori knowledge of the surface observed, the only knowledge exploited are those provided by images metadata.

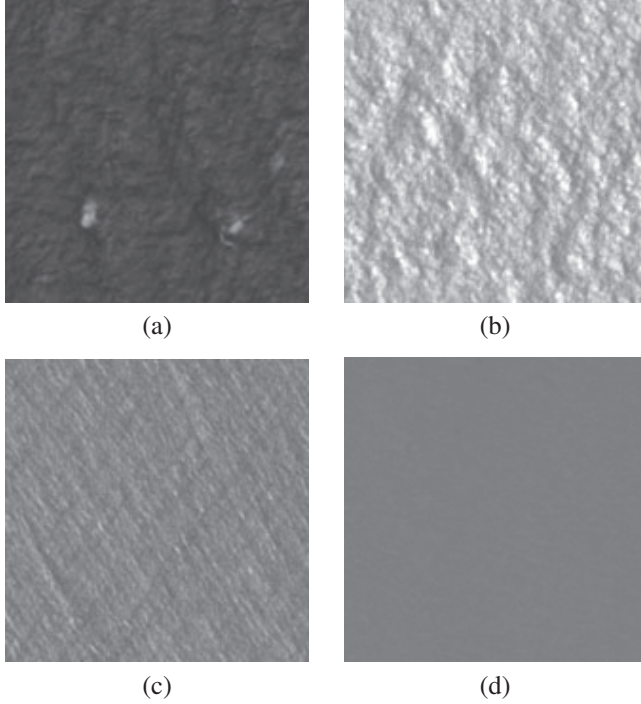
The paper layout is as follows. In section 2, we present sea clutter and its various components. We introduce the models to be tested in section 3 and present the used dataset in section 4. Fitting procedures and goodness-of-fit test are introduced in section 5, the obtained results in section 6 and conclusion in section 7.

## 2. SEA CLUTTER ON OPTICAL IMAGES

The sea clutter on high-resolution optical images is the result of the integration of multiple reflectances related to phenomena such as capillary waves, swell and whitecaps. In those images, the sea-surface may appear very differently according to its state and acquisition geometry and illumination. It can be seen both uniform, textured or studded with whitecaps as shown in Figure 1 thereby producing very different histograms from Gaussian like distribution to heavy tail ones.

Considering a homogeneous sea surface, sea clutter can be divided into several components: on the one hand a rough surface with capillary waves and swell, on the other the foam, spray and with whitecaps. These two components are induced by the wind, whitecaps starting to appear from 3 Beaufort and spray from 5 Beaufort. Most studies of the reflectance of the sea surface are limited to the study of rough surface. In this case the global radiation incident on the sea surface must be considered first. From this global radiation we could distinguish the direct radiation, the sunlight, and the diffuse radiation, the skylight [2]. The reflectance of the sea surface

We would like to thanks ASTRUM GEO-Information Services for funding the project ISAFUM (Imagerie Satellite et FUsion Multi-capteurs pour la surveillance maritime) in which this study is conducted and for providing satellite images.



**Fig. 1.** Example of thumbnails from which the histograms are extracted FORMOSAT2. Different sea states are visible: (a) fully developed sea with whitecaps, (b) textured sea with saturation maybe produced by high reflection or whitecaps, (c) cross swell, (d) calm sea. Images ©NSPO - 2011, Distribution Spot Image S.A.

is usually modeled simply in terms of parameters related to its roughness, as the distribution of slopes of the surface proposed by Cox and Munk [3, 4], under known conditions of acquisition and illumination. Here we propose the inclusion of all components of sea clutter in its statistical modeling.

### 3. TESTED STATISTICAL MODELS

As said above in 1 a lot of statistical model were tested for the detection of ship on SAR images. Crisp [5] make an exhaustive review of models used on this type of data. Gaussian distributions were used initially before switching to heavy-tailed distributions such as Weibull, Rayleigh and Gamma distribution. The latter allowed more efficiently to describe the non-Gaussian behavior of wind driven water surface waves observed by Onorato [6]. One of the most widely used and appropriate distributions for ship detection on SAR data is the K-distribution [7]. More recently the Alpha-stable distribution [8] has been used effectively for this purpose. We are particularly interested in it because of these characteristics detailed below and the Joelson's proposal [9] to use it to model slopes of a sea surface distribution in place of the Grams-Charlier series proposed by Cox and Munk.

Alpha-stable distribution is a four-parameter heavy-tailed distribution. This distribution is widely used as model of large events of random variations including extreme events. Zolotarev [10] proposed the following characteristic function,  $\alpha S_{(\beta, \gamma, \delta)}$ , describing Alpha-stable distribution:

$$\begin{aligned} & \text{if } \alpha \neq 1 \\ \alpha S_{(\beta, \gamma, \delta)}(t) &= e^{(it\delta - |\gamma t|^\alpha [1 + i\beta \tan(\frac{\pi\alpha}{2}) s(t) (|t|^{1-\alpha} - 1)])} \\ & \text{if } \alpha = 1 \end{aligned} \quad (1)$$

$$\alpha S_{(\beta, \gamma, \delta)}(t) = e^{(it\delta - |\gamma t| [1 + i\beta \frac{2}{\pi} s(t) \log |t|])}$$

with the characteristic exponent  $\alpha \in ]0, 2]$ , the skewness parameter  $\beta \in [-1, 1]$ , the scale parameter  $\gamma \in \mathbb{R}^{+*}$  and the location parameter  $\delta \in \mathbb{R}$ . The Alpha-stable probability density function,  $f\alpha S_{(\beta, \gamma, \delta)}$ , is obtained by calculating the Fourier transform of its characteristic function (1):

$$f\alpha S_{(\beta, \gamma, \delta)}(x) = \int_{-\infty}^{\infty} \alpha S_{(\beta, \gamma, \delta)}(t) e^{(-itx)} dt \quad (2)$$

Finally six distributions were selected to be tested: Gaussian, Weibull, Rayleigh, Gamma, K and Alpha-stable distributions.

### 4. DATASET

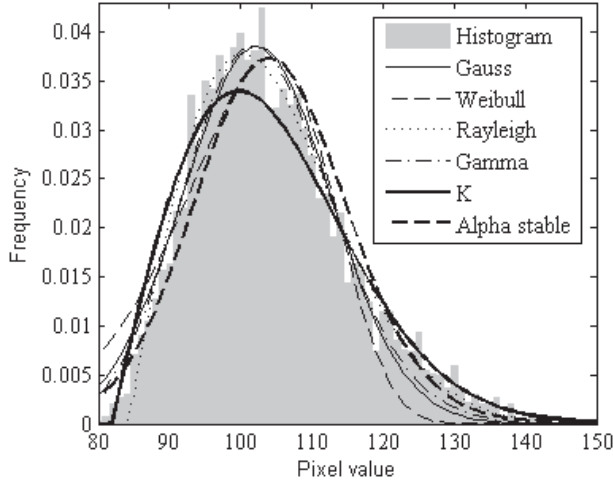
Images acquired by SPOT4, SPOT5, FORMOSAT2 and KOMPSAT2 were made available to this study by ASTRIUM GEO-Information Services. All images are panchromatic with pixel sizes of respectively 10, 5, 2 and 1 meter. They are radiometrically corrected of distortions due to differences in sensitivity of the elementary detectors of the viewing instrument. They are also geometrically corrected and georeferenced in a standard cartographic projection (UTM WGS84).

The less cloudy images were selected and among them we select a number of identical images for each satellite. This selection is done visually in order to cover a wide range of sea state. For a given image, variations in the surface were observed. We wish to extract histograms corresponding to homogeneous sea surfaces, therefore we subset images into 512 by 512 pixels thumbnails.

No atmospheric corrections were made. The images have been retained in numerical counts. The absolute value of the pixels is not essential in this study since we are interested in the distribution of these values into thumbnails. Also the histograms are normalized to facilitate their approximation by statistical models.

### 5. EXPERIMENTS

Dimensions 512 by 512 pixels were chosen because it allows the sampling of a homogeneous sea surface regardless of the sensor. However On spatial resolution in the other it is not the



**Fig. 2.** Histogram and distributions fit for a 64x64 pixels textured thumbnail

same wavelengths of waves that contribute to the inter-pixel or intra-pixel variability of the signal. The area covered by the thumbnail and the number of samples (pixels) contained in the thumbnail will also varied. Thus a study of the size of thumbnails best suited proves indispensable. In each initial thumbnail, thumbnails of six other dimensions 256x256, 128x128, 64x64, 32x32, 16x16, 8x8 were extracted. The histograms are calculated with a bin width equal to 1. Data range is reduced from 0 to 1 in order to fit properly probability density function not defined in a negative interval and to facilitate comparisons of different histograms. Finally the histograms are normalized for density estimation. All densities were estimated using the method of least squares. Figure 2 show an example of histogram and distribution fitting. To initialize the least-squares parameters of the distributions were first estimated by various methods: for Gaussian distribution mean and standard deviation are calculated, for Weibull, Rayleigh, and Gamma, distribution parameters are first estimate using the maximum likelihood method. Parameter estimation of K distribution is widely discussed in the literature. However many estimators produce outliers of preventing a good initialization parameters of least squares. The choice was made to initialize the scaling parameter by the minimum value of the data and the shape parameter by 1. Alpha-stable parameters are initialized using the Koutrouveliss iterative procedure [11].

Goodness-of-fit for all distribution and for all thumbnail sizes are evaluated by the Kolmogorov-Smirnov test and by the analysis of the sum of squared residuals provided by the least squared estimation.

Texture descriptors are calculated for each thumbnail: contrast, correlation, uniformity, homogeneity and entropy. Texture descriptors are calculated for each thumbnail: contrast, correlation, uniformity, homogeneity and entropy. A 2D Fourier transform is then used to extract the power spectral

density of thumbnails. Initially, the power spectral density is then averaged for each frequency regardless of the direction to obtain an indicator of predominant frequencies observed in the texture.

## 6. RESULTS

|              | 512  | 256  | 128  | 64   | 32   | 16   | 8    | All size |
|--------------|------|------|------|------|------|------|------|----------|
| Gauss        | 45.0 | 54.2 | 74.2 | 86.7 | 94.2 | 91.7 | 89.2 | 76.4     |
| Weibull      | 27.5 | 33.3 | 45.0 | 56.7 | 77.5 | 88.3 | 88.3 | 59.5     |
| Rayleigh     | 50.0 | 58.3 | 74.2 | 83.3 | 94.2 | 93.3 | 92.5 | 78.0     |
| Gamma        | 50.0 | 62.5 | 77.5 | 89.2 | 95.8 | 93.3 | 99.2 | 81.1     |
| K            | 64.2 | 72.5 | 80.0 | 81.7 | 87.5 | 95.8 | 96.7 | 82.6     |
| Alpha-stable | 90.0 | 93.3 | 96.7 | 99.2 | 97.5 | 93.3 | 90.8 | 94.4     |
| All models   | 54.4 | 62.4 | 74.6 | 82.8 | 91.1 | 92.6 | 92.8 | 78.7     |

**Fig. 3.** Success rate for the Kolmogorov-Smirnov test.

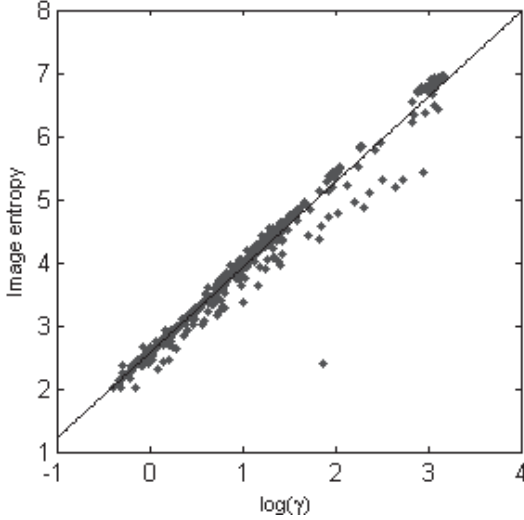
On table 3 all the distribution except Weibull show good result, higher than 75%, in the Kolmogorov-Smirnov test. Gaussian, Rayleigh, Gamma, K s get comparable results on the entire dataset. Only the alpha-stable distribution comes off the lot with 94.4% success rate.

For all distributions, except the alpha-stable we observe that the success rate at test increases with decreasing size of thumbnails. This decrease is responsible for the decrease in success rate over the entire dataset. Indeed the Weibull distribution with dimensions of 512x512 thumbnails has a success rate of 28% while with dimensions of 16x16 it reached 88%. Sizes 32x32, 16x16, 32x32, 8x8, 8x8 and 64x64 allow to obtain the best success rate respectively for Gaussian, Weibull, Rayleigh, Gamma, and K, Alpha-stable distributions. Only the alpha-stable distribution is not very sensitive to the size of thumbnails with a success rate between 90% for dimension of 512x512 and 98% for dimension of 64x64.

|              | 512     | 256     | 128     | 64      | 32      | 16      | 8       | All size |
|--------------|---------|---------|---------|---------|---------|---------|---------|----------|
| Gauss        | 8.8E-05 | 8.8E-05 | 8.3E-05 | 7.0E-05 | 9.4E-05 | 2.0E-04 | 7.2E-04 | 1.9E-04  |
| Weibull      | 2.3E-04 | 3.7E-04 | 3.2E-04 | 2.6E-04 | 2.9E-04 | 7.8E-03 | 1.8E-02 | 3.9E-03  |
| Rayleigh     | 2.8E-04 | 2.6E-04 | 2.9E-04 | 3.5E-04 | 4.0E-04 | 6.8E-04 | 7.2E-03 | 1.4E-03  |
| Gamma        | 9.0E-05 | 7.5E-05 | 7.6E-05 | 1.0E-04 | 1.2E-04 | 2.3E-04 | 6.7E-04 | 1.9E-04  |
| K            | 3.5E-03 | 3.6E-03 | 4.0E-03 | 3.9E-03 | 4.0E-03 | 3.4E-03 | 3.0E-03 | 3.6E-03  |
| Alpha-stable | 2.0E-04 | 1.3E-04 | 2.7E-04 | 1.9E-04 | 2.6E-04 | 8.9E-04 | 7.0E-04 | 3.8E-04  |
| All models   | 7.3E-04 | 7.5E-04 | 8.4E-04 | 8.2E-04 | 8.6E-04 | 2.2E-03 | 5.1E-03 | 1.6E-03  |

**Fig. 4.** Mean sum of squared residuals in least-squares sense for the tested thumbnails size and distributions.

By analyzing sum of squared residuals in least-squares sense presented on table 4, we can see that those obtained for the K distribution are much larger than other distributions. The K-distribution, despite a high success rate in



**Fig. 5.** Comparison of entropy measured on thumbnails and the scale parameter  $\gamma$  of the Alpha-stable distribution.

Kolmogorov-Smirnov, strongly underestimates the contribution of the mode of the histogram. We also see a significant increase in the residue for sizes 16x16 and 8x8. Three distributions are getting the lowest residues, Gauss, Gamma and Alpha-stable.

Due to its best result and less sensitivity Alpha-stable distribution is retain as the best model. Texture and frequency descriptors are then compared with its tree significant parameters: the characteristic exponent, the skewness parameter, the scale parameter to assess relations binders. A single link between the texture descriptors or frequency and the parameters of the alpha-stable distribution has been found: there is a relation connecting the entropy and the scale parameter  $\gamma$ . For entropy values below 2, the performance of the least squares approximation of the Gamma brings a high variability of values of  $\gamma$ . Beyond the value 2 (textured images) a logarithmic relationship linking the two variables appears (3) with a 0.96 correlation coefficient as seen in figure 5.

$$\epsilon_{(\gamma)} = 1.3469 \log(\gamma) + 2.5855 \quad (3)$$

with  $\epsilon$  the thumbnail entropy and  $\gamma$  the scale parameter of Alpha-stable distribution.

## 7. CONCLUSION

Alpha-stable distribution appears to be best suited to model histograms. It is less sensitive to the size of thumbnails, the image resolution as well as changes in sea state has three degrees of freedom are certainly no stranger to these performance compared to other heavy tail distributions. The best approximation is obtained for window sizes of 64x64 with a success rate of 99.2% for Kolmogorov-Smirnov test. From the perspective of developing a ship detector these dimensions

appear to be appropriate under the criteria of Wackerman [7] for a background window of constant false alarm ship detector. Good performance for the 512x512 thumbnail size and good relationship between the Alpha-stable scale parameter and image entropy gives us confident about the potential performance of the use of Alpha-stable distribution in a ship detector on optical images.

Further studies will be conducted on the model's robustness to noise and the relationship between environmental parameters and model parameters.

## 8. REFERENCES

- [1] N. Proia and V. Pagé, "Characterization of a bayesian ship detection method in optical satellite images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 7, no. 2, pp. 226–230, 2010.
- [2] N.G. Jerlov, *Marine optics*, vol. 14, Elsevier Science Ltd, 1976.
- [3] C. Cox and W. Munk, "Measurement of the roughness of the sea surface from photographs of the sun's glitter," *J Opt Soc Amer*, vol. 44, 1954.
- [4] C. Cox and W. Munk, "Statistics of the sea surface derived from sun glitter," *J. Mar. Res*, vol. 13, no. 2, pp. 198–227, 1954.
- [5] DJ Crisp, "The state-of-the-art in ship detection in synthetic aperture radar imagery," 2004.
- [6] M. Onorato, AR Osborne, M. Serio, L. Cavaleri, C. Brandini, and CT Stansberg, "Observation of strongly non-gaussian statistics for random sea surface gravity waves in wave flume experiments," *Physical Review E*, vol. 70, no. 6, pp. 067302, 2004.
- [7] C.C. Wackerman, K.S. Friedman, W.G. Pichel, P. Clemente-Colon, and Li Xiaofeng, "Automatic detection of ships in radarsat-1 sar imagery," *Canadian Journal of Remote Sensing*, vol. 27, no. 4, pp. 371–378, 2001.
- [8] C. Wang, M. Liao, and X. Li, "Ship detection in sar image based on the alpha-stable distribution," *Sensors*, vol. 8, no. 8, pp. 4948–4960, 2008.
- [9] M. Joelson and M.C. Neel, "On alpha stable distribution of wind driven water surface wave slope," *Arxiv preprint arXiv:0810.0455*, 2008.
- [10] V.M. Zolotarev, *One-dimensional stable distributions*, vol. 65, Amer Mathematical Society, 1986.
- [11] I.A. Koutrouvelis, "An iterative procedure for the estimation of the parameters of stable laws," *Communications in Statistics-Simulation and Computation*, vol. 10, no. 1, pp. 17–28, 1981.